10

15

Abstract of the disclosure

Method and apparatus for determining a set of large sequences from an electronic data base

The instant invention relates to a method of and an apparatus for determining a set of large sequences from an electronic data base comprising a set D = $\{d1, ..., dn\}$ of transactions di $(1 \le i \le n)$ in a computer system with an implemented query module, each of the large sequences on the set D of transactions di having a support value greater than or equal to a given support value S, each of the transactions di of the set D being a sequence of items of a record E = $\{e1, \ldots, em\}$ of items ej $(1 \le j \le m)$. A set Lk (k>2) of large sequences is determined from the set D of transactions, the large sequences of set Lk each comprising exactly k items of record E in a respective order R_{LK} , and an associated support value S_{LK} on the sequence D of transactions each being greater than or equal to the given support value S, and nothing but sequences comprising two of the large sequences of set Lk-1, as partly overlapping partial sequences, with the respective order R_{LK-1} , being taken into account in determining set Lk.